Claims

What is claimed is:

1. A method for identifying one or more mean items for a plurality of items, J, each of said items having at least one symbolic attribute, each of said symbolic attributes having at least one possible value, said method comprising the steps of:

computing a variance of said plurality of items, J, for each of said possible symbolic values, x_μ, for each of said symbolic attributes; and

selecting for each of said symbolic attributes at least one symbolic value, x_μ, that minimizes said variance as the mean symbolic value.

2. The method of claim 1, wherein said mean symbolic value for each of said symbolic attributes comprises said mean of said plurality of items.

3. The method of claim 1, wherein said mean symbolic value for each of said symbolic attributes comprises one or more hypothetical items.

4. The method of claim 1, further comprising the step of assigning a label to said plurality of items using at least one symbolic value from said at least one mean of said plurality of items.

5. The method of claim 1, wherein said plurality of items are a cluster including similar items.

6. The method of claim 1, wherein said items are programs.

7. The method of claim 1, wherein said items are content.

8. The method of claim 1, wherein said items are products.

9. The method of claim 1, wherein said step of computing a variance is performed as follows:

$$\text{Var}(J) = \sum_{i \in J} (x_i - x_\mu)^2$$

where J is a cluster of items from the same class, x_i is a symbolic feature value for item i, and x_μ is an attribute value from one of the items in J such that it minimizes said Var (J).

10. A method for characterizing a plurality of items, J, each of said items having at least one symbolic attribute, each of said symbolic attributes having at least one possible value, said method comprising the steps of:

computing a variance of said plurality of items, J, for each of said possible symbolic values, x_μ, for each of said symbolic attributes; and

characterizing said plurality of items, J, with at least one mean item by selecting for each of said symbolic attributes at least one symbolic value, x_μ, that minimizes said variance as the mean symbolic value.

11. The method of claim 10, wherein said mean symbolic value for each of said symbolic attributes comprises at least one mean of said plurality of items.

12. The method of claim 10, further comprising the step of assigning a label to said plurality of items using at least one symbolic value from said at least one mean item.

13. The method of claim 10, wherein said plurality of items are a cluster including similar items.

14. The method of claim 10, wherein said mean symbolic value for each of said symbolic attributes comprises one or more hypothetical items.

15. The method of claim 10, wherein said step of computing a variance is performed as follows:

$$\text{Var}(J) = \sum_{i \in J} (x_i - x_\mu)^2$$

where J is a cluster of items from the same class, x_i is a symbolic feature value for item i, and x_μ is an attribute value from one of the items in J such that it minimizes said Var (J).

16. A system for identifying one or more mean items for a plurality of items, J, each of said items having at least one symbolic attribute, each of said symbolic attributes having at least one possible value, said system comprising:

a memory for storing computer readable code; and

a processor operatively coupled to said memory, said processor configured to:

compute a variance of said plurality of items, J, for each of said possible symbolic values, x_μ, for each of said symbolic attributes; and

select for each of said symbolic attributes at least one symbolic value, x_μ, that minimizes said variance as the mean symbolic value.

17. The system of claim 16, wherein said mean symbolic value for each of said symbolic attributes comprises said mean of said plurality of items.

18. The system of claim 16, wherein said mean symbolic value for each of said symbolic attributes comprises one or more hypothetical items.

19. The system of claim 16, wherein said processor is further configured to assign a label to said plurality of items using at least one symbolic value from said at least one mean of said plurality of items.

20. The system of claim 16, wherein said plurality of items are a cluster including similar items.

21. The system of claim 16, wherein said processor computes said variance as follows:

$$\text{Var}(J) = \sum_{i \in J} (x_i - x_\mu)^2$$

where J is a cluster of items from the same class, x_i is a symbolic feature value for item i, and x_μ is an attribute value from one of the items in J such that it minimizes said Var (J).

22. An article of manufacture for identifying one or more mean items for a plurality of items, J, each of said items having at least one symbolic attribute, each of said symbolic attributes having at least one possible value, comprising:

a computer readable medium having computer readable code means embodied thereon, said computer readable program code means comprising:

a step to compute a variance of said plurality of items, J, for each of said possible symbolic values, x_μ, for each of said symbolic attributes; and

a step to select for each of said symbolic attributes at least one symbolic value, x_μ, that minimizes said variance as the mean symbolic value.

23. A system for identifying one or more mean items for a plurality of items, J, each of said items having at least one

symbolic attribute, each of said symbolic attributes having at least one possible value, said system comprising:

means for computing a variance of said plurality of items, J, for each of said possible symbolic values, x_μ, for each of said symbolic attributes; and

means for selecting for each of said symbolic attributes at least one symbolic value, x_μ, that minimizes said variance as the mean symbolic value.